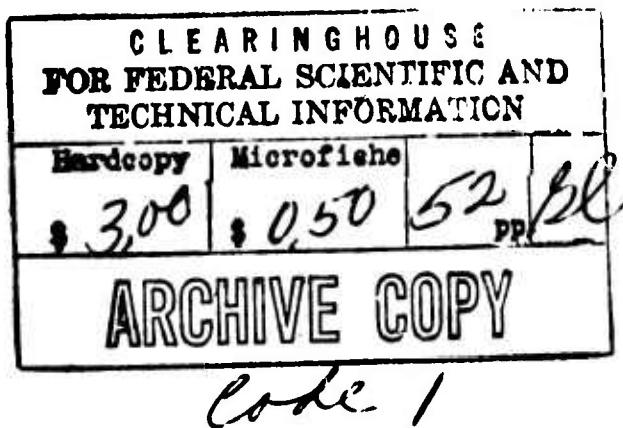


MEMORANDUM
RM-4443-PR
APRIL 1966

SOME STATISTICAL METHODS FOR MAINTENANCE ANALYSIS

Anders Sweetland



PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM

RM 4443-PR

APRIL 1966

**SOME STATISTICAL METHODS
FOR MAINTENANCE ANALYSIS**

Anders Sweetland

This research is sponsored by the United States Air Force under Project RAND—Contract No. AF 49(638)-1700—monitored by the Directorate of Operational Requirements and Development Plans, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

The RAND Corporation

1100 MAIN ST. • SANTA MONICA • CALIFORNIA • 90406

PREFACE

This Memorandum is written to supplement the statistical training of maintenance personnel in the Reports and Analysis section at the base level. (It should not be regarded as an introductory text in statistics.) Since it is intended as an adjunct to self-tutoring, it is written in language that the average NCO of the analysis section should find familiar and comprehensible. At critical points, statistical notation is explained so the reader can supplement his background by studying standard texts in statistics.

While making preliminary investigations for RAND research concerned with maintenance management, the author worked with the 15th Air Force in the analysis of the SAC Full-Force Project. It was discovered that the Reports and Analysis (R&A) personnel had not been exposed to certain statistical methods that are crucial in answering a number of questions. These methods, involving the use of chi-square and analysis of variance, are not presently taught in the R&A training program, even though they will handle a majority of the statistical problems usually arising at base level. It was further discovered that R&A personnel could readily learn these methods and just as readily put them to excellent use; their deep knowledge of the maintenance system made the tutoring process easy.

Incorporating the suggestions of several participants, this study formalizes the tutoring effort that resulted from the Full-Force experience. The Memorandum amounts to a simple translation of an elementary statistics text into the context of weapon-system analysis. Although most of the examples are taken from SACR 66-7 information, the methods are equally applicable to data generated by other major air commands.

This study, then, should be of interest to all personnel responsible for evaluating the effects of maintenance at any air base. It should also prove useful in resolving many of the analysis problems that arise at headquarters level.

SUMMARY

This Memorandum explains the use of two statistical tools, chi-square and analysis of variance, in the context of maintenance problems that arise at air bases.

These tools are particularly useful for examining the familiar day-to-day, month-to-month variation in maintenance numbers, the crucial question being whether given variations are the result of normal random fluctuation, or whether they are abnormal -- in which case they probably merit the expenditure of maintenance resources for corrective action.

Since most maintenance measures are frequency counts, the emphasis is on methods for dealing with this kind of data. Thus, following a brief introduction, the Memorandum asks a typical maintenance question: "How is alert affecting my break-rate?" The answers show how the chi-square test may be used. The method is elaborated with several examples.

Following this, two other maintenance questions demonstrate how simple analysis of variance may be used: "What effect is the 100-hour periodic having on subsequent missions?" and, "Do we have any bomb-nav systems going sour?" This Section includes several suggestions for easing the burden of computation. These suggestions are expanded in Appendix A, which shows how to use the PCAM to reduce the clerical effort.

ACKNOWLEDGMENTS

I am grateful to Ernest M. Scheuer and Sidney H. Miller of The RAND Corporation for a number of helpful suggestions, emendations, and additions to the manuscript. I wish also to thank Captain Joseph Cerny, SAC, who first introduced me to the problems of maintenance analysis.

The table of F, from George W. Snedecor's Statistical Methods (1956), is reproduced by permission of the author and the University of Iowa Press. For permission to reproduce the chi-square table from Statistical Tables for Biological, Agricultural, and Medical Research (1943), I am indebted to the authors, Sir Ronald A. Fisher and Dr. Frank Yates, and to the publishers, Oliver and Boyd, Ltd., of Edinburgh.

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGMENTS	vii
 Section	
I. INTRODUCTION	1
II. CHI-SQUARE: THE ONE-WAY CASE	2
III. CHI-SQUARE: THE TWO-WAY CASE	20
IV. SIMPLE ANALYSIS OF VARIANCE	26
 Appendix	
A. COMPUTATIONAL TECHNIQUES	37
B. SOME CRITICAL VALUES OF F AND CHI-SQUARE	41

I. INTRODUCTION

Only one thing is certain about maintenance numbers: they vary. Regardless of what the break-rate was last month, this month it will be either up or down. This variation is characteristic of stochastic phenomena, and it is at the heart of our problem. We know the numbers will change as a part of the normal fluctuation (i.e., random variation). What we need is a method for determining whether the change has exceeded the bounds of normal fluctuation. As one Deputy Commander for Materiel put it: "I know there has been a change; what I want to know is, should I worry about it?" How to answer his question is the objective of this study.

For example, our first illustration will be concerned with determining whether a two-week stay in ground alert has had a deteriorating effect on an aircraft. The numbers show that a number of first sorties after alert indicated an increase in the number of Form 126 write-ups as contrasted with other sorties. The question is: Should we attribute this increase to normal variation, or to the effects of the stay in alert? If the increase is only random variation, we will not want to waste manpower on special measures. Corrective action is the proper course if (and only if) the difference is significant -- that is, abnormal. Statistical techniques enable us to answer that question and thus conserve our maintenance resources for dealing with abnormal variation.

The following introduces Air Force maintenance personnel to chi-square and analysis of variance, and shows how these methods can be helpful in running the maintenance show.

II. CHI-SQUARE: THE ONE-WAY CASE

Consider the following questions: "How does a long period of alert status affect an aircraft?" "Is alert degrading the potential EWO (Emergency War Order) effectiveness?" "If so, what systems, if any, are most vulnerable?" "Can any steps be taken to prevent the degradation?"

To begin answering these questions and, incidentally, to begin learning a little about the use of chi-square (χ^2 , pronounced "ky-square"), the following numbers were collected:

1. Two samples of the number of write-ups on SAC 126 forms of 39 and 43 B-47 training sorties (204 and 221 write-ups, respectively).
2. A third sample of 124 write-ups produced by 18 sorties flown immediately after the aircraft came off alert.
3. Three similar samples produced by rolling dice. If we use dice, we can be sure that the resulting variation is due to randomness. (In order to get numbers resembling the actual maintenance data, three dice were used and three was subtracted from each throw. Our highest throw was 16 so the highest number in the table is 13.)

The numbers themselves (Table 1) do not show us any differences, either between samples or between a column of sortie write-ups and the adjacent column of dice rolls. If they were not labeled, we could not tell which set was which. Therefore, the conventional thing to do is to find the means (simple averages) of the numbers. Table 2 lists the average write-ups per sortie for the samples shown in Table 1.

In Table 2, we have attempted to capture the essence of this Memorandum's message. Both sets of samples show differing means. We know the differences are random in the dice-produced samples. With the aircraft samples we are not sure. Although the post-alert sample differs from the two regular-flyer samples, we do not know whether the difference is within normal bounds. Our quandary is only heightened by Table 3, which regroups the data in Table 1 to

Table 1

A COMPARISON OF ACTUAL WRITE-UPS OF TRAINING SORTIES
WITH "WRITE-UPS" PRODUCED BY DICE ROLLS

Number of Write-Ups					
Sample 1		Sample 2		Sample 3	
43 Dice Rolls	43 Regular Training Sorties	39 Dice Rolls	39 Regular Training Sorties	18 Dice Rolls	18 Post-Alert Training Sorties
13	12	12	13	11	13
11	11	11	9	11	12
10	10	10	9	11	12
9	10	9	8	10	11
9	10	9	8	10	9
9	8	9	8	7	8
8	8	9	8	7	7
8	7	9	7	6	7
8	7	9	7	5	7
7	7	8	7	5	7
7	7	8	7	5	6
7	7	8	7	4	6
7	7	7	7	4	4
6	7	7	6	4	4
6	7	7	6	4	3
6	6	6	6	2	3
6	6	6	5	1	3
6	6	6	5	0	2
5	6	6	5		
5	5	5	5		
5	5	5	5		
5	5	5	4		
5	5	5	4		
5	5	4	4		
5	4	4	4		
4	4	4	4		
4	4	4	4		
4	4	4	4		
4	3	4	4		
4	3	4	3		
4	3	3	3		
3	3	3	3		
2	3	3	2		
2	2	3	2		
2	2	3	2		
2	2	2	1		
2	1	2	1		
1	1	1	1		
1	1	1	1		
0	0	0	0		

Table 2

ARITHMETIC MEANS OF THE DATA
SHOWN IN TABLE 1

Sample	Write-Ups/Sortie	
	Dice	Aircraft
1	5.21	5.14
2	5.90	5.23
3	5.94	6.89

show the frequency with which sorties resulted in 13 write-ups, 12, 11, etc., down to 0. Thus Table 3 suggests that it would be prudent to "wait one" before going gung-ho on the assumption that alert

Table 3

FREQUENCY DISTRIBUTIONS OF THE DATA IN TABLE 1

Number of Write-ups	Frequency					
	Dice Sample			Aircraft Sample		
	1	2	3	1	2	3
13	1				1	1
12		1		1		2
11	1	1	3	1		1
10	1	1	2	3		
9	3	6			2	1
8	3	3		2	4	1
7	4	3	2	8	6	4
6	6	5	1	5	4	2
5	7	4	3	5	5	
4	6	7	4	4	8	2
3	1	5		5	3	3
2	5	2	1	3	3	1
1	3	1	1	4	3	
0	2		1	2		
Σ	224	230	107	221	204	124
N	43	39	18	43	39	18
M	5.21	5.90	5.94	5.14	5.23	6.89

NOTE: Column for Aircraft Sample 3 is read: one sortie yielded 13 write-ups, 2 sorties yielded 12, one sortie yielded 11, etc. Σ = total write-ups; N = number of sorties; M = mean.

aircraft are yielding more write-ups. While it is true that aircraft just off alert show more write-ups, on an average, than do the regular flyers, this difference apparently is all but lost in the variability: the distributions of the flyers scatter just like the distributions of the dice (i.e., from 0 to 13 write-ups). The question arises: "Is the scatter among the flyers random (as is the scatter of the dice), or is there a difference: is there reason to presume that the scatter among the flyers may not be random?"

The chi-square test is one method of checking for randomness, i.e., deciding whether differences among flyers can be attributed to chance. The following calculations are for the dice samples (where we know the answer):

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Base Line ("Sorties")	F_o	F_t	$F_o - F_t$	$(F_o - F_t)^2$	$\frac{(F_o - F_t)^2}{F_t}$
1	43	224	241.2	-17.2	295.8	1.23
2	39	230	218.8	+11.2	125.4	0.57
3	18	107	101.0	+ 6.0	36.0	0.36
Total	100	561	561.0			$\chi^2 = 2.16$

NOTE: F_o stands for observed frequency, F_t for theoretical frequency. χ^2 = chi-square = 2.16.

Details of the calculations follow. The first 43 "sorties" produced 224 write-ups out of the total of 561. Theoretically, since there were 100 sorties, we would have expected the 43 sorties to produce $(43/100)x(561) = 241.2$ write-ups. Similarly:

$$(39/100)x(561) = 218.8$$

$$(18/100)x(561) = 101.0$$

If the computation is correct, the sums of the observed (F_o) and theoretical frequencies (F_t) will agree within rounding error (561 and 561.0).

Column 5 tabulates the difference between F_o and F_t : $224 - 241.2 = -17.2$. This difference is squared in Col. 6: $(-17.2)^2 = 295.8$, which is then divided by F_t to give Col. 7: $295.8/241.2 = 1.23$. All the items in Col. 7 are added to get the chi-square: $1.23 + 0.57 + 0.36 = 2.16$.

To interpret the significance of the chi-square we need one additional bit of information: the degrees of freedom (df). For the one-sample chi-square method, the df is always one less than the number of categories. In this instance, since we are working with three categories (samples),

$$df = 3 - 1 = 2.$$

We then enter a table of significance levels for the chi-square distribution, at the row for $df = 2$:

Critical Values of Chi-Square

df	Probability that $\chi^2 \geq$ Chi-Square											
	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	etc.
1
2	1.39	2.41	5.99	...
etc.												

NOTE: See chi-square tables, Appendix B. Dots here indicate data that are irrelevant for our immediate use. Each line differs according to df, and in our case of 2 df the second line in the table applies.

The numbers across the top of the table (.99, . . . , .50, . . .) are called probability levels. These are the probabilities that a chi-square will be higher than the critical value shown down its column at the applicable df line. The probability level which is selected for a particular chi-square test is called the level of significance; all computed chi-square values greater than this critical value fall into the region of rejection. For example, the critical

value for the .05 significance level is 5.99, and all computed chi-square values for 2 df higher than 5.99 fall into the region of rejection for this particular chi-square test.

Our chi-square, 2.16, is between the critical values associated with the .50 and .30 probability levels for 2 df. This means that of all computed chi-squares based on sets of rolls of dice (where each set consists of 43, 39, and 18 rolls respectively) more than 30 to 50 of every 100 such sets would show a higher computed chi-square value. We can therefore say that chance accounts for the variation we found, since it is customary to reject the hypothesis of homogeneity -- i.e., random variation -- only if the computed value of χ^2 falls beyond the .05 significance level for the appropriate df. This choice of .05 is arbitrary; .01 could have been used and often is.

To recapitulate:

1. We collected some data: three samples of "write-ups."
2. We determined a theoretical distribution by apportioning our observations in terms of sorties flown.
3. We worked some simple arithmetic based on the formula:

$$\chi^2 = \sum \frac{(F_o - F_t)^2}{F_t}$$

(The Greek capital sigma (Σ) indicates a summation, i.e., that the numbers following should be added up.)

4. Finally, we checked the significance of our chi-square with a table of chi-square probabilities to see if the numbers were getting out of the range of normal variation.

Now that we have confirmed that the variations in the three samples of dice results are, in fact, random, let us look at the B-47 data of Table 1.

Samples	Sorties	Write-Ups	F_t	$F_o - F_t$	$(F_o - F_t)^2$	$\frac{(F_o - F_t)^2}{F_t}$
1	43	221	236.1	-15.1	228	0.96
2	39	204	214.1	-10.1	102	0.48
3	18	124	98.8	+25.2	635	6.42
Total	100	549	549.0			$\chi^2 = 7.86$ $df = 2$

NOTE: The computations are the same as before:

$$\begin{aligned}43/100 \times 549 &= 236.1 \\39/100 \times 549 &= 214.1 \\18/100 \times 549 &= 98.8\end{aligned}$$

Entering the χ^2 table, we find (for 2 df):

df	Probabilities		
	.05	.02	.01
1
2	5.99	7.82	9.21
3

There are only 2 chances in 100 of getting a χ^2 as large as we get (7.86). Consequently, we strongly suspect that something other than chance has been at work. Checking the computation, we find that the biggest contributor to the χ^2 is the third (Alert) sample and that the observed frequency is much greater than the expected (or theoretical) frequency. The aircraft are affected (the dice are "loaded"), apparently, by their stay in alert.

Before we consider a more elegant problem, some comment on chi-square analysis is in order.

1. Because the numbers are easy to get (simple sorting and tabulation), and the computations simple, chi-square is an exceedingly economical tool.
2. Chi-square does not "prove" or disprove anything. It only indicates whether the numbers can be explained by chance or not. Note: the chi-square did not prove that Alert status caused malfunctions; it only indicated that the differences among the numbers were probably not due to chance. The difference could result from dirty data or faulty sampling; it might even be that flight crews are more critical about aircraft just off alert and therefore write up more things. What chi-square has done, rather, is suggest a profitable area for further exploration (i.e., the effects of standing in alert). This can be of considerable importance to the short-staffed maintenance group, who must limit their efforts to areas with potential payoff.
3. One cannot get into too much trouble, if one remembers
 - a. The sum of F_o and F_t must agree (within rounding errors).

- b. The F_t should be greater than 5 in each cell (preferably greater than 10). (This is a requirement to meet certain mathematical conditions on which this procedure is based).
 - c. That chi-square only interprets the data given to it.
4. In the statistics books the formula is generally written:

$$\chi^2 = \sum \frac{(F_o - F_t)^2}{F_t},$$

and also:

$$\chi^2 = \sum \frac{(o-e)^2}{e},$$

in which "e" represents "expected or theoretical frequency", and "o" represents "observed frequency."

By way of review of the chi-square exercise, let us take another conventional maintenance question: Do we have any "dogs" or "hangar queens"? From last week's maintenance data collection (MDC) reports, we isolate the information shown in the table below:

Tail Number	Work Units	Man-Hours
248	6	9
135	16	28
426	21	28
186	22	27
261	32	48
285	36	58
403	37	52
187	51	96
210	58	106

From the looks of this, the big troublemakers are Tail Numbers 187 and 210, while the friends of maintenance are 248, 135, 426, and 186. But let us look at these same numbers statistically and in terms of numbers of sorties, as we did before. Here is the chi-square computation for work units:

Tail Number	Sorties Flown	Work-Units		$\frac{(F_o - F_t)^2}{F_t}$	Means (Observed Work- Units/Sortie)
		F_o	F_t		
248	2	6	16.91	7.04	3.00 (-)
135	1	16	8.45	6.73	16.00
426	3	21	25.36	0.75	7.00 (-)
186	3	22	25.36	0.45	7.33 (-)
261	5	32	42.27	2.50	6.40 (-)
285	4	36	33.82	0.14	9.00
403	4	37	33.82	0.30	9.25
187	3	51	25.36	25.91	17.00
210	8	58	67.64	1.37	7.25 (-)
Total	33	279	278.99	$\chi^2 = 45.19$	GM = 8.45
				df = 8	

NOTE: The minus signs in the "Means" column tag those aircraft that are below the fleet grand mean of 8.45 observed work-units per sortie flown.

Examination of the "critical values of chi-square" table, Appendix B, shows that our χ^2 of 45.19 for 8 df is far beyond the .001 level. This certainly is not random variation. A close examination of the computations reveals an entirely different picture from that in the table of MDC data above. Tail Number 210, rather than being the "dog," turns out to be a "good bird." As a matter of fact, it has produced a below-average number of work units per sortie flown. The real problem is Tail Number 187. Besides producing double the average number of work units, its contribution to the chi-square summation is far beyond bounds: the .05 level for 8 df is 15.51 and No. 187 alone contributes 25.36. Number '35 is the other aircraft that is far above the mean. We will want to keep an eye on it: if it keeps up at the rate it is going, it will be in the same class as No. 187.

The examples thus far have been simplified to ease the learning process. Let us look now at a typical real-life computation.

The samples of write-ups used for the first illustration were taken from a larger sample obtained from one of the SAC-47 bases (15th AF). The complete set of data and computations are:

Fly Cycle Number	Sortie Count	F_o (Write-Ups)	F_t	D^2/F_t	Mean Write-Ups per Sortie
1st	43	221	235.25	0.86	5.14(-)
2d	45	262	246.20	1.01	5.82
3d	46	210	251.67	6.90	4.57(-)
4th	39	204	213.37	0.41	5.23(-)
5th	29	143	158.66	1.55	4.90(-)
6th	23	135	125.83	0.67	5.87
AA	18	124	98.48	6.61	6.89
AP	16	118	87.54	10.60	7.38
$\Sigma = 259$		$\Sigma = 1417$	1417.00	χ^2 28.62	GM= 5.47
				df = 7	

NOTE: The cycle numbers refer to the monthly flying cycles; i.e., the 1st, 2d, 3d time an aircraft flew during a month. AA is the first sortie flown after alert. AP is the first sortie after a periodic. The AA and AP counts are not included in the other counts. D^2/F_t represents the differences (between F_o and F_t) squared, divided by F_t .

The table of chi-squares (see the 7 df row) shows that the computed value of χ^2 is beyond the 0.001 significance level. Thus, we reject the hypothesis that the differences among the eight measures are due to chance variation: there is less than one chance in 1000 of getting a chi-square this large.

The mean write-ups per sortie are included to further the comparison (e.g., the first sample consisted of 43 sorties, producing 221 write-ups, for an average of $221/43 = 5.14$ write-ups per sortie). The entire study involved 259 sorties producing 1417 write-ups for a grand mean of 5.47 ($1417/259 = 5.47$). Again, the minus signs tag those sample means that are smaller than the grand mean.

Checking the χ^2 column, we note that the 3d cycle, AA, and AP are the biggest contributors to the chi-square. Cross-checking with the Means column, we find that the mean for the 3d cycle is smaller than the grand mean, and those for AA and AP are much larger. Thus, we get a clearer picture of the numbers: the chi-square indicated that the difference among the eight samples was not due to chance; and by cross-checking with the chi-square and Means columns, we begin to get some hunches about which sets of data are abnormal. We have

an unanswered question: "Is something peculiar about Cycle 3?" It produces far fewer write-ups than expected. We would like to know what caused that.

Note well, however, that the chi-square test told us only that the differences among all eight samples were due to something other than chance. We cannot say, specifically, that Cycles 3, AA, and AP are the loaded dice, but they are good candidates for a more detailed investigation.

Our conclusion, based on the test, is that there is strong reason to suspect that post-alert and post-periodic sorties result in more write-ups than do the regular flyers. This is an interesting finding, but an equally interesting one would be the answer to the question: "Do aircraft just off alert have discrepancies that cause more frequent loss of mission than do the regular flyers?" We could dig into the SAC form 127 and the SAC debriefing forms file for two sets of data (the "0" and "1" code of Col. 14, which would yield a sample of regular flyers and AA's). Then we could sum the training items scheduled (Cols. 17-21) and the training items lost due to malfunction (Cols. 41-42), and perform the test almost as before.

Before we can calculate the chi-square for these data, we must learn one more thing. Until now we have obtained frequencies of write-ups, work units, etc., for each sortie, and then summed all the frequencies to get the total of frequencies in the category. In the next few examples, we will discuss an interesting, unique case where each observation can fall into one of two classes -- success or failure -- and we will get the observed and theoretical frequency in each class. Data where each observation is grouped according to success or failure, and the total frequency of successes and failures for all observations is listed, are called binomial distribution data. When the data are binomially distributed (as they will be in the next few examples) the observed and theoretical frequencies of both successes and failures must be included in the chi-square computations, as will be shown. Binomial distributions frequently encountered are:

Heads or tails;
Yes or no;
Broken or unbroken;
Write-ups or no write-ups;
Zero or one;
Failed or passed;
Successful or unsuccessful;
and so on.

The computation for binomially distributed data is illustrated.

Note that it requires only slight further computation.

First lay out and compute as before:

Class	Items Scheduled	F_o Successful	F_t Successful	D^2/F_t
Regular	697	633	622.6	0.17
Post-Alert	62	45	55.4	1.95
	759	678	678.0	

NOTE: As before, $678/759 \times 697 = 622.6$, etc.

Then arrive at the nonoccurrence $F_o + F_t$ by subtraction:
 $697 - 633 = 64$, and $62 - 45 = 17$.

Class	Items Scheduled	F_o Unsuccessful	F_t Unsuccessful	D^2/F_t
Regular	697	64	74.4	1.45
Post-Alert	62	17	6.6	16.39
	759	81	81.0	

Combine D^2/F_t .

$$\begin{aligned}0.17 + 1.45 &= 1.62 \\1.95 + 16.39 &= 18.34 \\2 \\x &= 19.96 \\df &= 1\end{aligned}$$

The chi-square is highly significant: a check of the chi-square table shows that it is well beyond the 0.001 level. Our alert aircraft are contributing an undue share (17 items lost versus an expected 6.6).

Let us suppose that the findings did not come out so neatly. Assume that χ^2 came out 1.72 (instead of 19.96). A check with the chi-square tables (for 1 df) would show that the probability of getting a chi-square as large as or larger than that observed lies between 0.20 and 0.10. In other words, there are about 15 chances in 100 of getting a χ^2 this big or bigger -- that is, of having observations differ from theoretical values by as much as or more than actually occurred even if regular and post-alert situations are the same; and therefore the differences occurred by chance. This probability (0.15) is in the messy, in-between area. One cannot say that it is low enough (i.e., $p = 0.05$) to reject a null hypothesis (meaning, in this case, that the difference is due to chance). But, even though the direction is the same as that of the previous data, we cannot be sure there is a significant difference. Several alternatives are possible, however:

First, we can assume there is nothing sacred about an 0.05 or 0.01 probability level, which is a matter of convention. If the situation is critical -- for example, if it might affect a unit's showing in an Operational Readiness Inspection (ORI) -- we probably would take steps to see that things did not get worse.

Second, if the matter is not critical, we might elect to keep it under close surveillance, checking to see if the data-trend continues to deteriorate.

Third, we might obtain an entirely different set of numbers (say from last quarter, or from some other base) and run it again.

While we are sorting the SAC Form 127, we would probably wish to pick up data on three other critical measures to determine whether a stay in alert affects:

1. The number of late takeoffs due to materiel.
2. The number of cancellations due to materiel.
3. The EWO effectiveness.

Each of these is a critical measure that might be influenced by standing in home alert. The chi-square testing would be identical to that in the previous example: two categories (1 df) using the "0" and "1" codes. Since these are all binomial distributions, the non-occurrence items must be included in the computation.

The examples have been taken from the 1962 SAC 126 form; the new SAC form works out even better for SAC analyses. Late takeoffs, cancellations, first sorties of the day, first sorties after ground alert, and other data (Cols. 17-31) are processed as described. In addition, the information from Cols. 32-51 provides an excellent set of data to monitor for possible "sick" systems. The Code 1 counts (system used and satisfactory) become the Base-Line, and the 2, 3, and 4 codes (system used and not satisfactory), combined, become the observed frequencies (i.e., write-ups).

The new ADC mechanized debriefing form (76-3) lends itself equally well to statistical analysis. The sortie cards yield the write-up counts (by type of mission if desired) as well as verification of armament connections. However, to get intercept effectiveness, some minor PCAM operation is necessary. The easiest way is to punch the tail number in Cols. 78-80 on all cards. This enables one to isolate data by sortie and serial number. With this, some really elegant analyses are available: which tail-numbers have out-of-bounds kill rates under what circumstances; what situations produce the worst and the best kill ratios; what flight crews need special training in which situations, and so on.

One of the advantages of using chi-square analysis is that it provides a means of directing the resources of R&A to payoff areas. Thousands of such investigations could be undertaken by R&A, but many would end up as complete wastes of effort. With chi-square, we can make a quick, cheap test to determine if a more elegant investigation will be profitable. The following is a partial list of things that might be looked at:

- Comparison of experienced and inexperienced flight crews
- Comparison of experienced and inexperienced maintenance personnel
- The effects of long versus short training missions
- First 25 hours versus last 25 hours since the 100-hour periodic
- Yo-Yo versus non-Yo-Yo flights
- Low-altitude versus high-altitude sorties
- Reciprocating versus jet engines
- Winter versus summer

Regular sorties versus special sorties
Morning versus evening flyers.

Most of these are general, however. Let us look at something more specific: getting leverage on the SAC and TAC Management and Control System (MCS) data. (Note that we have already touched on several MCS measures: late takeoffs, cancellations, training loss, and materiel defects.)

The testing of MCS data is exactly the same as before; in addition, we will keep pace with progress by using time-sampling techniques. The intent is to catch any odious trend before it is too late.

MCS scoring is comprised of two parts: the Base-Line and a related measure (i.e., the observed frequency):

1. Manning in Required Specialties:

Base-Line: Total Requirement
 F_o : Total Assigned
(Officers and Airmen kept separate)

2. Individual Proficiency Training:

Base-Line: Total Eligible
 F_o : Total in Training
Base-Line: Number Testing
 F_o : Number Passing
etc.

The chi-square test can also be used to monitor for out-of-bounds conditions. In practice, the monitoring agency (R&A) establishes a suitable time-sample (every week, two weeks, month, etc.) and determines the number of samples to process, (e.g., four consecutive two-week samples).

For example, assume that Shop Reparable Performance is to be monitored. It is decided that four consecutive two-week samples will be used in the computations. The Base-Line is Items Processed. The F_o is Items Repaired. These are binomially distributed data, so the nonoccurrence (nonrepaired) counts must be included. The computations are:

Month	Base-Line	F_o	F_t	$\frac{D^2}{F_t}$	Means
July	811	721	665.60	25.72	0.889
	941	762	772.29	0.76	(-)0.810
August	763	628	626.20	0.03	0.823
	932	718	764.91	$\chi^2 = 16.04$	(-)0.770
		2829	2829.00	$\chi^2 = 42.55$ df=3	GM=0.821

Note: $\frac{D^2}{F_t}$ column includes occurrences and nonoccurrences.

In the example, the observed chi-square was 42.55, which is well beyond the .001 significance level for χ^2 with 3 df, and the last sample in August showed an unfortunate downward trend. Since this is MCS, it is critical. We want to act.

Assume that steps were taken to improve the shop repairable performance. The method of handling the resulting data is shown below. The first two weeks data are dropped, and the new data are added on the end. The chi-square is recomputed:

Month	Base-Line	F_o	F_t	$\frac{D^2}{F_t}$	Means
July	941	762	760.25	0.02	(-)0.810
August	763	628	616.44	1.13	0.823
September	932	718	752.98	8.46	(-)0.770
	826	689	667.34	$\chi^2 = 3.66$ df=3	0.834 GM=0.808
		2797	2797.01		

The numbers have now taken a trend back up. The September data show Shop Repairable to have a "batting average" of 0.834, while the over-all "batting average" is 0.808 (the Grand Mean).

With the August data removed, the numbers show only conventional fluctuations of random events; hence, repairable performance is in a steady-state condition. Note that chi-square does not (and cannot) say whether an over-all batting average of 0.808 is good enough for the base repairable performance.

The second method of monitoring MCS data has already been implied: determining the effects of some corrective program. Data are divided into two categories, "Before" and "After". The question to be asked is: "Did the new maintenance program significantly improve performance?"

Remember that any data can show a change for the better, but this change may be only a random fluctuation -- that is, no real change at all. The DCM needs to know whether the change is significant or only random variation (as seen in the last example).

The rationale of the previous paragraph follows:

- a. If the upward trend is only a random fluctuation, the DCM needs to know this so that further measures can be made to insure that the trend is truly upward.
- b. Or, if the upward trend is only random variation, and the cost of the new procedures is great, the DCM may wish to go back to the old procedures that gave the same results at lower cost. More likely, he may wish to revise the procedures again. It is worthwhile to repeat the point: the mere movement of up or down does not necessarily mean either improvement or regression; the movement may be due to random fluctuation. The basic question is: "Is the upturn outside the limits of chance--has there been a true improvement?"

An example of testing for improvement follows: Inadequate Shop Reparable performance provoked the establishment of a special set of procedures to improve performance. The "Before" and "After" data are tested. The Base-Line is total items processed, and F_o is total items repaired.

Time	Base-Line	F_o	F_t	D^2/F_t	Means
Before	1695	1346	1366.03	1.51	(-)0.794
After	<u>1752</u> <u>3447</u>	<u>1432</u> <u>2778</u>	<u>1411.97</u> <u>2778.00</u>	<u>1.46</u> $\chi^2 = 2.98$ df=1	<u>0.817</u> GM=0.806

Here is a good illustration of how chi-square can be helpful. Apparently the new set of procedures has yielded an improvement: the number of items processed is up (1346 to 1432), and the "batting average" is up (0.79 to 0.82). But, as the chi-square shows, there is no cause for optimism: the probability is approximately 0.10. We cannot say certainly that the new set of procedures has had any effect. Some re-evaluation of the situation is in order.

This concludes the discussion of the use of chi-square with the one-way case. Before going on to the two-way case, we may profitably repeat certain points:

1. Numbers vary as part of the normal, random fluctuation. Consequently, they can make false impressions. Chi-square minimizes this false impression by providing a test for homogeneity.
2. The chi-square test is easily computed and hence is cheap to use. (Most of the examples used in this discussion were extracted from regular records.)
3. A check should always be made to insure that the sums of the F_o and F_t agree within rounding errors, and that the F_t are greater than 5.0 in each cell.
4. If the data are in the form of binomial distributions, the nonoccurrence items must be included in the computation.
5. The data must be in the form of frequency counts.
6. Chi-square tests only the numbers offered to it, not the reasoning behind the selection of the numbers.

III. CHI-SQUARE: THE TWO-WAY CASE

The one-way use of the chi-square was characterized by use of a Base-Line and one set of observations (the F_0). The two-way case uses no Base-Line, and two or more sets of observed frequencies. The absence of a Base-Line restricts the possible applications, for reasons to be discussed later. But where it can be used, the two-way case is an elegant means of testing a broad array of data to find out if something is getting out of balance.

For illustration, let us test the classic belief that maintenance troubles come in cycles: "This month, it is Bomb-Nav, last month it was ..."

As before, a chi-square is computed and checked to see what probability level it represents. The computational process is slightly different. The observations (write-ups) are first set up as shown in Table 5.

Table 5
ORGANIZATION OF WORKSHEET FOR TWO-WAY COMPUTATIONS

System	Write-Ups			
	July	August	September	Row Total
Communication	272	243	231	746
Navigation	208	225	218	651
ECM	286	302	316	904
Bomb-Nav	407	388	411	1206
Autopilot	68	76	55	199
Column Total	1241	1234	1231	3706

In computing, we first sum both the rows and columns to get the margin totals. As a check, we get the matrix total twice, by summing both the row and the column margin totals.

The theoretical frequency (F_t) for each cell is computed:

$$F_t = (\text{Col. Total}) \times (\text{Row Total}) / (\text{Matrix Total});$$

thus the July Bomb-Nav F_t is

$$F_t = (1241) \times (1206) / (3706) = 403.84.$$

With a desk calculator, a lot of the tedium can be circumvented by first obtaining (for each column) the constant:

$$k = (\text{Col. Total}) / (\text{Matrix Total}),$$

and then multiplying each Row Total by the constant.

The chi-square as before is:

$$\chi^2 = \sum \frac{(F_o - F_t)^2}{F_t}.$$

In each cell, F_t is subtracted from the corresponding F_o ; the difference is squared and divided by F_t . As before, these results are summed to get chi-square.

A complete picture of what has happened can be obtained by recording three entries in each cell: the F_o , the F_t , and the D^2/F_t . In addition, when the χ^2 is significantly different from chance, one should locate those cells with large D^2/F_t and put a "+" or "-" in front of the F_o that are bigger or smaller than the F_t . This gives us some "hunch material" for use when we start digging deeper.

Table 6 is the completed computation of the previous data. The degree-of-freedom computation is:

$$df = (R - 1) \times (C - 1).$$

In Table 6, we have 5 rows and 3 columns; hence:

$$df = (5 - 1) \times (3 - 1) = 8.$$

In this example, the χ^2 is at the probability level:

$$.50 > p > .30.$$

Table 6
COMPLETE COMPUTATION FOR DATA GIVEN IN TABLE 5

System	Entry	July	August	September
Communication	F_o	272.00	(-)243.00	(-)231.00
	F_t	249.81	248.40	247.79
	D^2/F_t	1.97	0.12	1.14
Navigation	F_o	(-)208.00	225.00	218.00
	F_t	218.00	216.76	216.24
	D^2/F_t	0.46	0.31	0.01
ECM	F_o	(-)286.00	302.00	316.00
	F_t	302.72	301.01	300.28
	D^2/F_t	0.92	0.00	. 0.82
Bomb/Nav	F_o	407.00	(-)388.00	411.00
	F_t	403.84	401.57	400.59
	D^2/F_t	0.02	0.46	0.27
Autopilot	F_o	68.00	76.00	(-)55.00
	F_t	66.64	66.26	66.10
	D^2/F_t	0.03	1.43	1.86
Chi-square = 9.82				
df = 8				

Hence, we conclude that the system is showing the normal variation of a steady-state condition. We would single out no system for special attention.

Now assume that the computation has been set as a continuous monitoring function and that the October results were those of Table 7.

Table 7
COMPUTATIONS FOR AUGUST-OCTOBER DATA

System	Entry	August	September	October
Communication	F_o	(-)243.00	(-)231.00	285.00
	F_t	248.11	247.50	263.39
	D^2/F_t	0.10	1.10	1.77
Navigation	F_o	225.00	218.00	(-)212.00
	F_t	214.11	213.59	227.30
	D^2/F_t	0.55	0.09	1.03
ECM	F_o	302.00	316.00	(-)296.00
	F_t	298.78	298.05	317.18
	D^2/F_t	0.03	1.08	1.41
Bomb/Nav	F_o	(-)388.00	411.00	(-)419.00
	F_t	398.15	397.18	422.67
	D^2/F_t	0.26	0.48	0.03
Autopilot	F_o	76.00	(-) 55.00	98.00
	F_t	74.86	74.68	79.47
	D^2/F_t	0.02	5.19	4.32
$\text{Chi-square} = 17.46$ $df = 8$				

The observed χ^2 of 17.46 falls between the .05 and .02 levels of the tables of the χ^2 distribution for 8 df. It is unlikely that the distribution of the F_o is due to random variation. Perusal of the matrix suggests that something is peculiar with the Autopilot maintenance. It was way down last month, and way up this month. We have some "hunch material" to pursue.

Note: The test did not show that Autopilot was the culprit. The test indicated that the entire matrix was not homogeneous. The only thing we know is that the system (as represented by the five measures) has slipped out of a steady-state condition. We have a hunch -- and it is only a hunch -- that the Autopilot data are peculiar.

Hence, we have some means of logically ordering our attention to find reasons why.

In the two-way case, it is particularly desirable to insure that F_t in each cell is greater than 5 and preferably greater than 10. There will be little trouble with AFM 66-1 data, but debriefing data will give some problems. However, there is a way out: one can combine rows or columns where such combination will not render the results meaningless.

For example, "notorious troublemakers" were used in the previous example. Had we desired to include some other systems (such as 41000, 45000, 47000, 49000), it probably would not have been possible to get sufficient F_t to compute. In this instance, all these entries might be combined into one row: "Miscellaneous utilities, hydraulic, and pneumatic systems."

Let us sketch a few matrices that might be set up as monitoring devices. We shall use a three-month time sample to make the discussion easier, but this time span is not sacred, since the matrix is not limited to three columns. We can use any number from two on up. Nor is a month's sample mandatory. It could be lesser or greater (i.e., weekly, biweekly, quarterly). But weekly samples of much 66-1 data are liable to be subject to erratic collection procedures (e.g., sometimes the data are in by Friday afternoon and sometimes not). This will clobber the χ^2 , which will be high and erratic. As before, the nonoccurrence contribution must be included where binomial data are used.

Some examples follow:

Suggested Categories	Time Sample		
	July	August	September
Suggested Variables			
Basic aircraft			
Power plants			
Utilities			
Instruments and autopilot			
Communications			
Weapon delivery			
	Write-ups		
	Units Produced		
	Etc.		

We would also use the type of sortie flown instead of a time sample to determine if there is a differential effect. For example, if it is believed that high-altitude sorties put increased demand on power plants and electronic equipment, our categories could be high, medium, and low altitude.

Various combinations are testable. For example, there is some reason to believe that the type of countermeasures combined with the tracking angle have a differential effect on intercept success. This could be determined.

Countermeasures	Track Angle		
	Frontal	Beam	Stern
Chaff	Write-ups		
ECM	Missed intercepts,		
None	ground environment error		
Both	Missed intercepts, material		
	Missed intercepts, flight crew		

In general, it will be more desirable to use the one-way case for measuring MCS data, such as SACR 66-7 and ADCR 66-28, thereby taking advantage of the Base-Line used by that method.

This completes the discussion of the two-way case. Two cautions should be repeated:

1. For mathematical reasons, the theoretical cell frequencies should be greater than 5, and preferably greater than 10. The small theoretical frequency problem can occasionally be circumvented by combining categories (e.g., two months' data instead of one, or combining several similar categories into an "others" or a "miscellany" category).
2. The chi-square tests the entire matrix for nonrandom variation. The method of isolating large contributors to the chi-square is only a method to help infer logical areas to explore with a more detailed investigation.

IV. SIMPLE ANALYSIS OF VARIANCE

In Sec. I the phrases "more detailed investigation," "more careful study," "more critical testing" appeared frequently. The following discussion treats one of these methods in detail: analysis of variance.

The heart of analysis of variance is the famous F-test, in which:

$$F = \frac{\text{between-groups variance}}{\text{within-groups variance}}.$$

Having computed F, we enter an F-table. (Unlike chi-square, the F-test involves entering two numbers (df) instead of one to get our probability level.) The process is simple, but attempts to explain the rationale behind it frequently end in confusion. Rather, let us run through the mechanics of computation.

The example is similar to the one used before, consisting of 3 samples: 222 write-ups produced by 43 regular training sorties, 204 write-ups produced by 39 regular training sorties, and 300 write-ups produced by 38 training sorties immediately after the aircraft came out of a hundred-hour periodic (see Table 8). The question is: "Do periodics cause additional work?"

The tedious part of the computation is getting the sum of the d^2 (Σd^2). We need the sum of the raw scores (ΣX) and the sum of the raw scores squared (ΣX^2):

$$\Sigma d^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

It is convenient to do this by sample, as in Table 9.

Table 8

WRITE-UPS PRODUCED BY TWO SAMPLES OF 43 AND 39 REGULAR TRAINING SORTIES AND ONE SAMPLE OF 38 TRAINING SORTIES WITH AN AIRCRAFT JUST OUT OF PERIODIC

Sample 1	Sample 2	Sample 3
12	13	13
11	9	12
10	9	12
10	8	11
10	8	11
8	8	10
8	8	10
7	7	10
7	7	10
7	7	10
7	7	9
7	7	9
7	6	9
7	6	9
6	6	9
6	6	8
6	5	8
6	5	8
5	5	8
5	5	8
5	4	8
5	4	8
5	4	7
4	4	7
4	4	7
4	4	7
4	4	7
3	4	7
3	3	6
3	3	6
3	3	4
3	2	4
2	2	3
2	2	3
2	1	3
1	1	2
1		
1		
1		
0		

Table 9
ILLUSTRATIONS OF BASIC COMPUTATIONS
FOR GETTING THE Σd^2

Scores (Sample 1)

X	X^2
12	144
11	121
10	100
10	100
10	100
8	64
Etc.	
1	1
1	1
1	1
1	1
1	1
0	0
$\Sigma X = 222$	$\Sigma X^2 = 1516$

Once the ΣX and the ΣX^2 are computed, 90 percent of the work is done.

Two techniques are available to reduce the burden:

1. With a desk calculator, the ΣX and the ΣX^2 are obtained in one pass by squaring each X using the "accumulating-multiply" button. (It took less than 2 minutes to get the complete ΣX and ΣX^2 for Sample 1, illustrated in Table 4.)
2. A simple EAM procedure will give ΣX , ΣX^2 , and N. See Appendix B.

In a similar manner, we get ΣX and ΣX^2 for each of the samples:

	Sample 1	Sample 2	Sample 3	Total
ΣX	222	204	300	726
ΣX^2	1516	1320	2626	5462
N	43	39	38	120

The basic computations follow.

We are trying to get two numbers: 1) Σd^2 between samples; and 2) Σd^2 within samples. The easiest way to get these two numbers is to first get the Total Σd^2 and then subtract the between Σd^2 . The process follows:

To get the total $\sum d^2$, add up the $\sum X$ and $\sum X^2$ obtained (Total column -- see preceding tabulation) and apply the formula

$$\sum d^2 = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$\begin{aligned} \text{Total } \sum d^2 &= 5462 - \frac{(726)^2}{120} \\ &= 5462 - 4392.3 \\ &= 1069.70. \end{aligned}$$

The between $\sum d^2$ is somewhat easier. The individual $\sum X$'s and n's are used:

$$\text{Between } \sum d^2 = \frac{(222)^2}{43} + \frac{(204)^2}{39} + \frac{(300)^2}{38} - \frac{(726)^2}{120} = 189.34.$$

Note that the correction term $\frac{(726)^2}{120}$ is the same in both equations.

The within $\sum d^2$ is obtained by subtraction:

$$\begin{aligned} \text{Within } \sum d^2 &= \text{total } \sum d^2 - \text{between } \sum d^2 \\ &= 1069.70 - 189.34 = 880.36 \end{aligned}$$

The last needs are the degrees of freedom. As before, df is always something minus one. In this instance:

$$\text{Total df} = 120 - 1 = 119 \text{ (120 measures)}$$

$$\text{Between df} = 3 - 1 = 2 \text{ (3 samples)}$$

$$\text{Within df} = 119 - 2 = 117 \text{ (total df - between df)}$$

Or one can derive the within df directly (it is a good check) by going directly to the sample n's.

$$\begin{aligned} 43 - 1 &= 42 \\ 39 - 1 &= 38 \\ 38 - 1 &= 37 \\ \text{Total} &= 117 = \text{within df} \end{aligned}$$

Table 10 has all the information laid out in conventional format. (Although the total entries are not used, they are included to give the complete story.)

Table 10

COMPLETE ANALYSIS-OF-VARIANCE SUMMARY

Source of Variation	Sum of Squares	df	Mean Square	F
Between	189.34	2	94.67	12.58
Within	880.36	117	7.52	
Total	1069.70	119		

The mean square is obtained by dividing each $\sum d^2$ by its df. ($189.34/2 = 94.67$ and $880.36/117 = 7.52$). $F = 94.67/7.52 = 12.58$.

Reading the table of the F distribution (Appendix B) is slightly different from reading a table of chi-square. In reading F-tables one must find the cell represented by the two df's. (In this example, between df = 2 and within df = 117, so we go to the second column and 117th row. There is no row 117, so we take row 125, the next best thing.) The cell in the F table shows two entries: one for the .05 (the smaller number) and one for the .01 levels of probability. We check to see if ours is bigger than either of these.

We might digress for a moment to explain a confusing element in the way F-tables are conventionally laid out. The common tables have a note, "degrees of freedom (for greater mean square)" over the columns. Strike the word "greater" and insert the word "between". You must use rows to find within, and columns to find between. You will quickly discover that if the computed F is less than 1.0, there is no point in bothering to look it up.

In our example, $F = 12.58$ and the 2×125 cell contained 3.07 and 4.78. Our F is bigger than 4.78 (the .01 level), so we say: the odds are more than 100 to 1 against our getting an F this large by chance. We reject the hypothesis that the differences among the samples is due to random variation. If our F had been smaller than 3.07 (the .05 level), we would have said that there are more than 5 chances in 100 of getting an F this big just from random variation. Therefore, we could not say that the numbers have any significance to us -- that is, random variation could account for any of the differences. (Or,

to use previous terms, the numbers are the same as we might have obtained by rolling the same set of dice.)

But the F test of the sample was high ($p < .01$). Something, apparently, has happened to cause the numbers to vary as they did. The numbers suddenly become interesting. Let us compute the means. We already have the numbers:

$$\text{Mean} = \Sigma X / N$$

$$\begin{aligned}\text{Regular Flyer 1} & \dots 222/43 = 5.16 \\ \text{Regular Flyer 2} & \dots 204/39 = 5.23 \\ \text{Post-periodic} & \dots 300/38 = 7.89\end{aligned}$$

The Post-periodic mean is obviously not what it should be. It is costing us an average of 2.7 additional write-ups. We will want to check the nature of the write-ups to see if preventive measures are possible.

This is, perhaps, too much discussion and not enough figuring. Let us take another critical problem. Records are kept of drops made by five different bomb-nav systems, measured in "yards from the shack". The question is: "Have one or more systems 'gone sick' or is the variation among the five systems random?"

The data and computations follow:

Systems				
(1)	(2)	(3)	(4)	(5)
89	07	38	27	97
04	86	31	32	98
98	83	91	66	51
41	08	38	11	40
28	41	02	17	26
65	00	11	62	63
05	86	48	62	63
39	33		75	36
64	61		87	93
65	43			29
	29			97
				85
				75

NOTE: The first drop for System 1 was 89 yards off, the second 04 yards off, etc.

	System					Total
	1	2	3	4	5	
ΣX	498	477	259	439	818	$\Sigma \Sigma X = 2,491$
ΣX^2	34,098	30,975	14,559	27,401	61,688	$\Sigma \Sigma X^2 = 168,721$
N	10	11	7	9	13	N = 50

NOTE: The expression $\Sigma \Sigma$ means "the sum of the sums."

$$\begin{aligned} \text{Total sum of squares} &= \Sigma X^2 - \frac{(\Sigma X)^2}{N} \\ &= 168,721 - \frac{(2,491)^2}{50} = 44,619.38. \end{aligned}$$

$$df \text{ total} = N - 1 = 50 - 1 = 49.$$

$$\begin{aligned} \text{Between sum of squares} &= \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \dots - \frac{(\Sigma \Sigma X)^2}{N} \\ &= \frac{(498)^2}{10} + \frac{(477)^2}{11} + \dots - \frac{(2,491)^2}{50} = 3,850.76. \end{aligned}$$

$$df \text{ between} = \text{number of samples} - 1 = 5 - 1 = 4.$$

$$\begin{aligned} \text{Within sum of squares} &= \text{total sum of squares} - \text{between sum of squares} \\ &= 44,619.38 - 3,850.76 = 40,768.62 \end{aligned}$$

$$\begin{aligned} df \text{ within} &= \text{total} - \text{between} \\ &= 49 - 4 = 45. \end{aligned}$$

Source of Variation	Sum of Squares	df	Mean	F
Between Groups	3,850.76	4	962.69	1.06
Within Groups	40,768.62	45	905.97	
Total	44,619.38	49		

Checking the F Table with the 4 and 45 df, we find the level of significance does not come close to the .05 level. This means that the between-systems deviation is the same variation we would get if we had drawn all the samples from a table of random numbers, which these were. And our Commander or the armament and electronics (A and E) squadron can relax with the assurance that none of the systems has gone sour.

One of the advantages of analysis of variance is nicely illustrated. As with chi-square, one can test a lot of samples simultaneously; we could easily have tested 10, 20, or 30 systems, without much increase in effort. (Had we tested each sample against each other sample by the conventional mean-difference methods, we would have had to compute 10 separate tests of mean differences, and still would not have the "big picture".) But with analysis of variance we do the testing in one fell swoop. The shortcoming of analysis of variance (like chi-square) is that the entire set of measures is tested; thus, we cannot (with analysis of variance alone) isolate the specific offenders. However, as we shall show, there is a way of solving this problem.

The analysis of variance is admirably suited for monitoring trends, if the suggested way of computing (i.e., by samples) is followed.

Assume that we are interested in monitoring the number of training items lost due to equipment malfunction. For each training sortie we record the number of items lost, and at the end of each week the sorties are counted (N), the number of items lost in each sortie is simultaneously summed (ΣX), and squared and summed (ΣX^2). We also compute the mean. The accumulations appear thus:

	Weeks			
	1	2	3	4
ΣX	11	8	12	12
ΣX^2	15	10	22	24
N	16	16	15	14
Mean	0.68	0.50	0.80	0.86

In practice, the data above are monitored. Any time the trend seems to be taking a turn for the worse (as in Week 4,) the analysis of variance is computed to determine if the variation is out of bounds.

Source of Variation	Sum of Squares	df	Mean Square	F
Between weeks	1.14	3	0.38	0.55
Within weeks	39.55	57	0.69	
Total	40.69	60		

Entering the F tables with 3 and 57 df, we find that the F does not reach the 0.05 level; hence, we infer that the variation shown is random.

It is possible to draw a picture of what the analysis of variance has shown us (i.e., whether or not the sample means are outside the limits of the mean of the entire set of samples).

To do this we need the grand mean of the set of samples and the standard error of the grand mean (σ_m). This is painless work since we have already done the drudgery. The problem used before will serve.

The totals obtained were:

$$\Sigma X = 726$$

$$\Sigma X^2 = 5462$$

$$N = 120$$

$$\Sigma d^2 = 1069.70 \text{ (i.e., } 5462 - \frac{(726)^2}{120} \text{)}.$$

The grand mean is:

$$726/120 = 6.05.$$

We now need to find the standard error of the mean:^{*}

$$\sigma_m = \sqrt{\frac{\Sigma d^2}{N(N-1)}}.$$

Hence the standard error of the mean is:

$$\sigma_m = \frac{1069.70}{(120)(119)} = 0.27.$$

The range included by the mean $\pm 2 \sigma_m$ includes 95 percent of the grand means we would get if we repeated the sampling hundreds of times.

^{*}The formula for the standard error of the mean can be found in any statistics textbook, which will show also that the sample mean has a normal or bell-shaped distribution with approximately 95 percent of the sample mean distributed about the true mean plus or minus two standard errors, σ_m .

In this case,

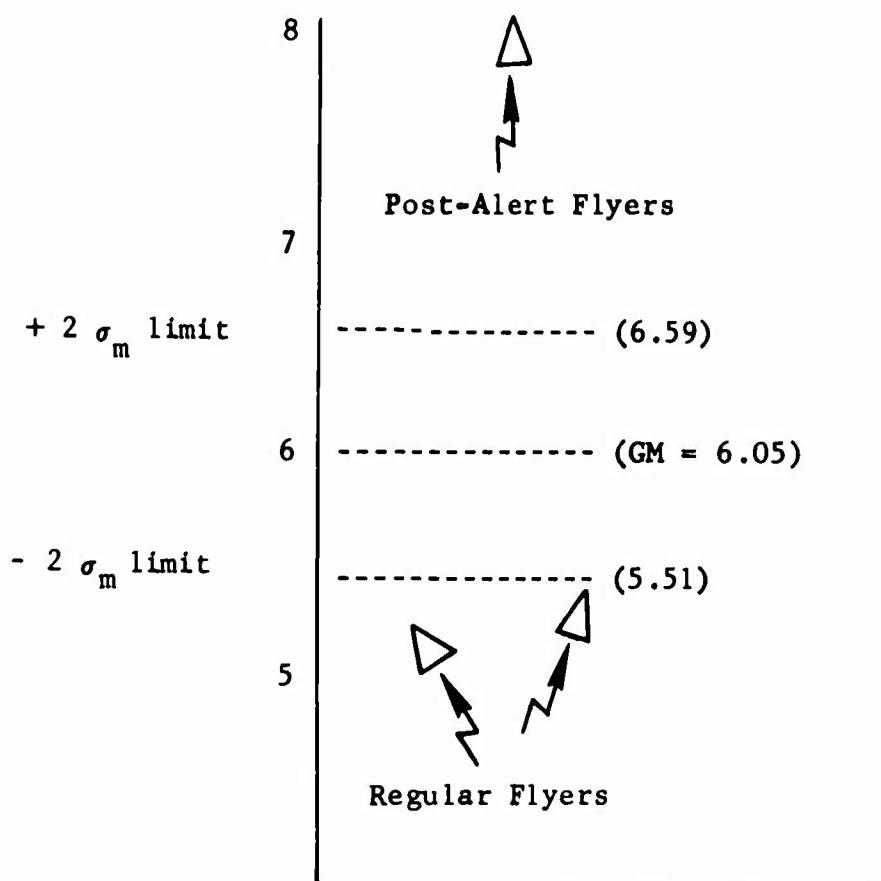
$$2 \sigma_m = (2)(0.27) = 0.54$$

$$\text{lower limit} = 6.05 - 0.54 = 5.51$$

$$\text{upper limit} = 6.05 + 0.54 = 6.59.$$

Of the sample means, 95 percent would fall between 5.51 and 6.59. Or 5 percent would fall outside the range 5.51 - 6.59. This is the same 5-percent level ($p = 0.05$) we have used before.

We can now make a picture of our results by plotting the means and the $\pm 2 \sigma_m$ limits. (The horizontal axis has no consequence.)



Note that grand mean (GM) provides a Base-Line. The critical factor is the distance from this Base-Line, while the $\pm 2\sigma_m$ limits provide meaning to this distance. In the example, we can see that the samples of regular flyers are considerably less than $1/2 \sigma_m$ apart, and both are several σ_m from the post-alert flyers.

The F-test told us there is little likelihood that all three samples were drawn from the same population. Plotting the data in terms of GM and σ_m gives us an idea of what has happened. We feel fairly sure that the two regular-flyer samples come from one population and that the post-alert flyers come from another. We now are fairly confident that a long stay in home alert does have a degrading effect on the aircraft.

Appendix A
COMPUTATIONAL TECHNIQUES

The most tedious part of any statistical computation is getting ΣX (sum of the raw scores) and ΣX^2 (sum of each raw score squared). Once these are obtained, the remaining calculations rarely take more than 15 minutes per problem.

The first suggestion is: never copy data. Record them in the form in which they will be used. In the case of AFM 66-1 and SACM 66-7 data, such a form is provided. For special reports, we suggest using AF Form 1530 (the 80-column key-punch form) for worksheets. Then, if the number gets too frequent for hand computation, the sheets can be key-punched as a preliminary to taking advantage of the Base PCAM.

There are three basic rules:

1. Never do a calculation by hand if you have a calculator (or slide rule).
2. Never do it by calculator if you have a table of squares.
3. Never follow either of the two previous rules if PCAM is available. Generally speaking, PCAM can add and subtract readily, but can multiply and divide only with extreme anguish.

SINGLE-SAMPLE CHI-SQUARE

If the data have been key-punched, obtain a listing like the following (have the PCAM group sum the columns):

Base-Line (BL) Data	F_0
8	7
4	4
6	5
<u>etc.</u>	<u>etc.</u>
$\Sigma = 128$	$\Sigma = 95$

The computation of F_t is:

Theoretical frequency = $L_1 / \Sigma BL \times \Sigma$ observed frequencies; e.g.,

$$X_{F_t} = (8/128) \times (95).$$

Since $95/128 = 0.742$ is a constant, lock 0.742 in the keyboard, then multiply each Base-Line entry by the constant.

Only two accuracy checks are possible. ΣF_o must equal ΣF_t (within rounding error) or the use of the chi-square will be invalidated. And the ΣD (differences between F_o and F_t) will equal zero (within rounding error).

Because the D^2 may become large, it is tempting to divide D^2 by a constant. If you do, do not forget to multiply it by the same constant afterward; otherwise the resulting chi-square will be shrunk to insignificance. Some people like to use the 10^n approach:

$$3216 = 3.216 \times 10^3$$

MULTI-SAMPLE CHI-SQUARE

If the data matrices are large (more than 50 cells) or if there are many of them, one should think about using PCAM for assistance. Without strain, PCAM can give both the sums of the rows (they call it cross-footing) and the sums of the columns (accumulating) with the grand total thrown in free.

Again the use of AF Form 1530 is urged. When you have many different kinds of samples to run, try to use the same columns on the Form 1530's. This way extensive rewiring of boards is avoided.

ANALYSIS OF VARIANCE

When the data come in strung out over time, you can avoid a flap by computing as you go along--that is, computing the ΣX and ΣX^2 at the end of each day. Then little work remains by the end of the week. Also, it is easier to check calculation accuracy when a large sample is broken into several small parts. (If you can get a printing calculator, accuracy checks are painless.)

The use of PCAM can be a big aid, but it takes a little prior planning and discussion with the PCAM people. There are several solutions. The one shown is not necessarily the best, but it is the easiest to understand.

Lay out the samples on the 1530 Form:

Sample 1	Sample 2	Sample 3
24	47	94
69	35	31
85	63	102
etc.	etc.	etc.

Then have key-punch make up a deck of χ^2 cards:

Card 1	1	1
Card 2	2	4
Card 3	3	9
Card 4	4	16
etc.		

Sample 1 of the data cards is sorted in the same order as the χ^2 cards. The cards are merged (on the collater) and the tab is wired to print both X and χ^2 for each X in the data. The card count is the N. (If zero is a significant measure--such as "no malfunctions" --include a zero card in the χ^2 deck to make the card-count correct.)

The data cards are then re-sorted on Sample 2 and the process repeated. If the data samples are on separate cards, you can also get minor and intermediate totals, along with card counts--in short, the complete information given in the tables used to illustrate the method.

Appendix B
SOME CRITICAL VALUES OF F AND CHI-SQUARE

Table 11
CRITICAL VALUES OF CHI-SQUARE

df	Probability under H_0 that $\chi^2 \geq$ chi square													
	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	.00016	.00063	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	.12	.18	.35	.58	1.00	1.42	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	.30	.43	.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	.87	1.13	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	3.57	4.18	5.23	6.30	7.81	9.03	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	5.23	5.98	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.46	23.54	26.30	29.63	32.00	39.29
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.75
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.60	22.76	25.99	28.87	32.35	31.80	42.31
19	7.63	8.57	10.12	11.65	13.72	15.35	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	8.90	9.92	11.59	13.24	15.44	17.18	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	9.54	10.60	12.34	14.04	16.31	18.10	21.24	24.94	27.30	30.81	33.92	37.66	40.29	48.27
23	10.20	11.29	13.09	14.85	17.19	19.02	22.34	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	10.86	11.99	13.85	15.66	18.08	19.94	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	11.52	12.70	14.61	16.47	18.94	20.87	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	12.20	13.41	15.38	17.29	19.82	21.79	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	12.88	14.12	16.15	18.11	20.70	22.72	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	13.56	14.85	16.93	18.94	21.59	23.65	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	14.20	15.57	17.71	19.77	22.48	24.58	28.34	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	14.95	16.31	18.49	20.60	23.36	25.51	29.34	33.53	36.25	40.26	43.77	47.96	50.89	59.70

SOURCE: Sir Ronald A. Fisher and Dr. Frank Yates, Statistical Tables for Biological, Agricultural, and Medical Research, Oliver and Boyd, Ltd., Edinburgh, 1943.

Table 12
VALUES OF F AT THE FIVE PER CENT (LIGHTFACE TYPE) AND ONE PER CENT (BOLODFACE TYPE) LEVELS OF SIGNIFICANCE

n ₁	n ₂ degrees of freedom (for greater mean square)																						
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500
1	4.052	4.999	5.463	5.625	5.764	5.859	5.928	5.981	6.032	6.082	6.136	6.142	6.169	6.208	6.234	6.253	6.286	6.313	6.334	6.352	6.361	6.366	2.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.48	19.49	19.49	19.50	19.50
3	98.49	99.89	99.17	99.25	99.30	99.33	99.34	99.36	99.38	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.49	99.49	99.49	99.50	99.50
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.64	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.73	2.72	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.21

* Reproduced from G. W. Snedecor. *Statistical Methods*. Ames, Iowa: Collegiate Press, 1940, pp. 184-187, by permission of the author and the publisher.

Table 12 -- Continued

n	n: degrees of freedom (for greater mean square)																								
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
14	4.60	3.74	3.34	3.11	2.91	2.85	2.77	2.70	2.65	2.60	2.55	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.14	2.13		
	5.36	6.51	5.36	5.63	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.79	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02	3.00	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07	
	5.32	6.26	5.42	4.89	4.56	4.32	4.14	4.04	3.99	3.89	3.73	3.67	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89	2.87	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01	
	5.23	6.23	5.29	4.77	4.44	4.26	4.03	3.89	3.78	3.69	3.61	3.55	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.89	2.86	2.80	2.77	2.75	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96	
	5.18	6.11	5.18	4.67	4.34	4.16	3.93	3.79	3.68	3.59	3.52	3.45	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67	2.65	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.27	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	
	5.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.89	1.88	
	5.18	5.93	5.61	4.59	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.19	3.12	3.09	3.00	2.92	2.84	2.76	2.70	2.63	2.60	2.54	2.51	2.49
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84	
	5.10	5.85	5.54	4.94	4.43	4.10	3.87	3.71	3.56	3.45	3.37	3.30	3.23	3.13	3.05	2.94	2.86	2.77	2.69	2.63	2.56	2.53	2.47	2.44	2.42
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.11	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81
	5.02	5.78	5.57	4.87	4.37	4.04	3.81	3.65	3.51	3.40	3.31	3.24	3.17	3.07	2.99	2.88	2.80	2.72	2.63	2.58	2.51	2.47	2.42	2.39	2.36
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78	
	5.04	5.72	5.52	4.82	4.31	4.01	3.76	3.59	3.45	3.35	3.26	3.18	3.12	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.37	2.33	2.31
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	
	5.04	5.64	5.44	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07	3.07	2.97	2.89	2.78	2.70	2.62	2.53	2.48	2.41	2.37	2.32	2.26
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.77	1.74	1.73	
	5.02	5.61	5.42	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.09	3.03	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.36	2.33	2.27	2.23	2.21
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71	
	5.00	5.57	5.37	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	3.05	2.99	2.91	2.79	2.62	2.54	2.45	2.40	2.33	2.29	2.23	2.19	2.17	2.16
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.86	1.82	1.78	1.76	1.72	1.70	1.69	
	4.73	5.53	5.33	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13

The function, $F = e$ with exponent $2s$, is computed in part from Fisher's table VI (7). Additional entries are by interpolation, mostly graphical.

Table 12 -- Continued

n	n ₁ degrees of freedom (for greater mean square)																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67
	7.63	5.49	4.66	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.96	2.93	2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.95	2.90	2.86	2.71	2.64	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.06	3.00	2.92	2.87	2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.90	2.84	2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59
	7.58	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.86	2.80	2.70	2.63	2.51	2.42	2.34	2.25	2.20	2.12	2.06	2.02	1.98	1.96
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57
	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.97	2.89	2.82	2.76	2.66	2.59	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56
	7.39	5.25	4.35	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.78	2.72	2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.87
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.98	1.92	1.88	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53
	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.75	2.69	2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.93	1.88	1.84
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.73	2.64	2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81	1.78
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49
	7.27	5.15	4.29	3.89	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.76
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.76	1.72	1.67	1.63	1.58	1.53	1.52	1.50	1.48
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.64	2.56	2.42	2.36	2.22	2.13	2.04	1.98	1.93	1.86	1.80	1.76	1.72
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45	
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58	2.48	2.40	2.32	2.20	2.11	2.02	1.96	1.91	1.84	1.78	1.73	1.70

Table 12 -- Continued

n ₁	n ₁ degrees of freedom (for greater mean square)																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.44	
	7.17	5.86	4.29	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.54	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68
65	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.83	1.81	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41
	7.12	5.81	4.16	3.68	3.37	3.15	3.06	2.98	2.85	2.75	2.66	2.59	2.53	2.43	2.35	2.23	2.15	2.06	1.96	1.92	1.82	1.78	1.71	1.66
80	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
	7.08	4.98	4.13	3.65	3.34	3.12	3.05	2.95	2.82	2.72	2.63	2.56	2.49	2.32	2.26	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37
	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.79	2.70	2.61	2.54	2.47	2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.75	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.51	2.45	2.35	2.26	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.47	1.42	1.39	1.34	1.30	1.28
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43
125	3.92	3.07	2.68	2.44	2.30	2.17	2.08	2.01	1.95	1.90	1.86	1.82	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.65	2.56	2.47	2.40	2.33	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.54	1.51	1.43	1.37	1.33
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19
	6.76	4.71	3.83	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.24	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.26
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.76	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.29	2.23	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.33	1.24	1.19
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08
	6.64	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.19	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.32	1.26	1.19
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.04	1.94	1.88	1.83	1.79	1.73	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.31	1.26	1.17	1.11	1.00
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.16	2.07	1.99	1.87	1.79	1.69	1.61	1.55	1.41	1.36	1.25	1.18	1.08

DOCUMENT CONTROL DATA

1. ORIGINATING ACTIVITY

THE RAND CORPORATION

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

SOME STATISTICAL METHODS FOR MAINTENANCE ANALYSIS

4. AUTHOR(S) (Last name, first name, initial)

Sweetland, Anders

5. REPORT DATE
April 19666a. TOTAL NO. OF PAGES
526b. NO. OF REFS.
---7. CONTRACT or GRANT NO.
AF 49(638)-17008. ORIGINATOR'S REPORT NO.
RM-4443-PR

9. AVAILABILITY/LIMITATION NOTICES

DDC 1

9b SPONSORING AGENCY

United States Air Force
Project RAND

10. ABSTRACT

A translation of an elementary statistics text into the context of weapon-system analysis. It is designed to supplement the statistical training of maintenance personnel of the Reports and Analysis section at the base level. Most of the examples are from SACR 66-7, but the methods are applicable to data generated by other major Commands. Since most maintenance measures are frequency counts, emphasis is on dealing with this kind of data. Three principal maintenance problems are considered: (1) how alert is affecting the break-rate; (2) what effect the 100-hour periodic is having on subsequent missions; and (3) whether any bomb-nav systems are deteriorating. The answers demonstrate the use of the chi-square test and simple analysis of variance.

11. KEY WORDS

Statistical methods & processes
Maintenance
Weapon systems